

Information filtering apparatus for filtering information for interests of users and a method therefor

Patent Number: ☐ US6014654
Publication date: 2000-01-11
Inventor(s): ARIYOSHI YUSUKE (JP)
Applicant(s): NIPPON ELECTRIC CO (JP)
Requested Patent: ☐ JP9265478
Application Number: US19970828060 19970328
Priority Number(s): JP19960075781 19960329
IPC Classification: G06F17/00
EC Classification: G06F17/60B2
Equivalents: JP2800769B2

Abstract

An information filtering apparatus estimates an evaluation value (suitability) of a demander with respect to an item which is not evaluated by the demander but is evaluated by users other than the demander. The information filtering apparatus estimates the suitability by studying relation among the demander and the other users with respect to an evaluation value evaluated for each item by the demander and the other users. A suitable estimating means may compute the suitability in consideration of the population density or an Leading-Edge degree. The suitability can be estimated by computing the Leading-Edge degree.

Data supplied from the **esp@cenet** database - I2

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平9-265478

(43)公開日 平成9年(1997)10月7日

(51)Int.Cl.⁶

識別記号

庁内整理番号

F I

技術表示箇所

G 0 6 F 17/30

G 0 6 F 15/403

3 5 0 C

3 4 0 A

審査請求 有 請求項の数5 OL (全 12 頁)

(21)出願番号

特願平8-75781

(22)出願日

平成8年(1996)3月29日

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 有▲吉▼ 勇介

東京都港区芝五丁目7番1号 日本電気株式会社内

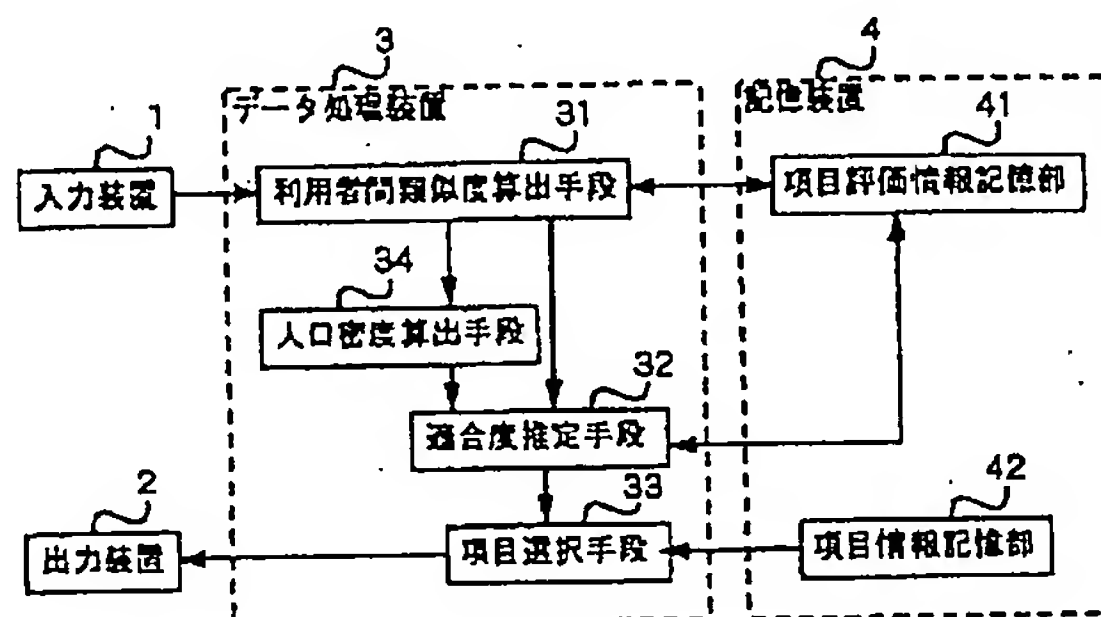
(74)代理人 弁理士 若林 忠

(54)【発明の名称】 情報フィルタリング方式

(57)【要約】

【課題】 フィルタリング精度の高い情報フィルタリング方式を提供する。

【解決手段】 多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、項目評価情報を使用してフィルタリング要求をした要求者 と他の利用者との類似度を算出する利用者間類似度算出手段と、項目評価情報および類似度から要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、適合度から情報やサービスを選択して出力する項目選択手段とを有する情報フィルタリング方式において、利用者間類似度算出手段の算出結果から似た興味や嗜好を持つ利用者の人数を求め、利用者近傍での局所的な人口密度を計算する人口密度算出手段を有し、適合度推定手段は、項目評価情報、類似度、及び人口密度から適合度を推定する。



【特許請求の範囲】

【請求項1】 多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、

推薦要求やフィルタリング要求が入力される入力装置と、

情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、

前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者との他の利用者との類似度を算出する利用者間類似度算出手段と、

前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、

該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、

前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、

前記利用者間類似度算出手段の算出結果から、似た興味や嗜好を持つ利用者の人数を求め、該人数から各利用者近傍での局所的な人口密度を計算する人口密度算出手段を有し、

前記適合度推定手段は、

前記項目評価情報、前記類似度、及び前記人口密度から前記適合度を推定することを特徴とする情報フィルタリング方式。

【請求項2】 多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、

推薦要求やフィルタリング要求が入力される入力装置と、

情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、

前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者との他の利用者との類似度を算出する利用者間類似度算出手段と、

前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、

該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、

前記記憶装置に、

前記項目評価情報とともに、情報やサービスに対する評価を行った時刻の情報を各利用者毎に蓄積し、

前記時刻情報から利用者の評価情報に対する信頼度の高さを示すLE度を算出するLE度算出手段を有し、

前記適合度推定手段は、

前記項目評価情報、前記類似度、及び前記LE度から前記適合度を推定することを特徴とする情報フィルタリング方式。

【請求項3】 多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、

推薦要求やフィルタリング要求が入力される入力装置と、

10 情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、

前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者との他の利用者との類似度を算出する利用者間類似度算出手段と、

フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する第1の適合度推定手段と、

該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、

20 前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記記憶装置に、

前記項目評価情報とともに、情報やサービスに対する評価を行った時刻の情報を各利用者毎に蓄積し、

前記項目評価情報および前記類似度から、各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す推定適合度を算出する第2の適合度推定手段と、前記推定適合度及び前記時刻情報から利用者の評価情報に対する信頼度の高さを示すLE度を算出するLE度算出手段とを有し、

30 前記第1の適合度推定手段は、

前記項目評価情報、前記類似度、及び前記LE度から前記適合度を推定することを特徴とする情報フィルタリング方式。

【請求項4】 多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、

推薦要求やフィルタリング要求が入力される入力装置と、

40 情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、

前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者との他の利用者との類似度を算出する利用者間類似度算出手段と、

前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、

該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、

50 前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、

前記項目評価情報とともに、情報やサービスに対する評価を行った時刻の情報を各利用者毎に蓄積し、

前記時刻情報から利用者の評価情報に対する信頼度の高さを示すLE度を算出するLE度算出手段を有し、

前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記記憶装置に、

適合度を推定する際に使用される各情報に対して、情報やサービスの分野及び利用者層に応じて付与される重み情報が蓄積される重み情報記憶部を有し、

前記適合度推定手段は、

前記項目評価情報、前記類似度、および前記重み情報から前記適合度を推定することを特徴とする情報フィルタリング方式。

【請求項5】 請求項4に記載の情報フィルタリング方式において、

前記重み情報は前記入力装置から入力されることを特徴とする情報フィルタリング方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は多数の情報やサービス中からの利用者の興味や嗜好に合う情報やサービスを推薦する情報フィルタリング方式に関するものである。

【0002】

【従来の技術】情報フィルタリング方式とはデータベース、パソコン通信、およびWWW (World Wide Web) などの情報システムが提供する大量の情報や多数のサービス（以下対象項目と称す）の中から、利用者の興味や嗜好に合うもの推薦したりフィルタリングする方式である。

【0003】従来の情報フィルタリング方式では、興味対象そのものである一次情報を直接検索して取り出す方式や、題名、著者名、キーワード、単語出現頻度などの二次情報を利用してフィルタリングを行ない所望の情報を取り出す方式が採用されている。

【0004】ところで、近年、対象項目に対する複数の利用者の評価結果を蓄積し、その項目評価情報を用いて利用者の興味や嗜好に合う項目を選び出し、推薦やフィルタリングを行なうソーシャルフィルタリングまたはコラボレーティブフィルタリングと呼ばれる方式が研究・開発されている。

【0005】このソーシャルフィルタリング方式の例として、例えば、プロシーディングス・オブ・ザ・シーエスシーダブリュ1994 (proceedings of the cscw 1994) 1994年エイシーエムプレス (acm press) 175頁から186頁米国では、MIT (Massachusetts Institute of Technology: マサチューセッツ工科大学) で行っている音楽用アルバムを対象にしたフィルタリングシステムについて紹介され、プロシーディングス・オブ・ザ・シーエイチアイ95 (proceedings of thechi 95) 1995年エイシーエムプレス (acm press) 210頁から217頁米国では、電子メールを対象としたフィルタリングシステムについて紹介されている。

【0006】しかしながら、各分野の情報量が飛躍的に

増大しているため、推薦やフィルタリング精度のさらなる向上が望まれている。

【0007】次に、従来のソーシャルフィルタリング方式について図面を参照して説明する。

【0008】図9はソーシャルフィルタリング方式の構成を示すブロック図であり、図10は図9に示した方式の動作手順を示すフローチャートである。

【0009】図9において、従来の方式は、推薦要求やフィルタリング要求（以下要求と称す）が入力される入力装置1と、要求した利用者（以下要求者と称す）に対する各対象項目毎の適合度を推定するデータ処理装置11と、複数の利用者による項目評価情報が蓄積される記憶装置12と、要求結果を表示する出力装置2とによって構成されている。

【0010】記憶装置12は項目評価情報記憶部121と項目情報記憶部122とによって構成され、項目評価情報記憶部121には複数の利用者による項目評価情報が蓄積され、項目情報記憶部122には対象項目を識別するためのID番号や項目名などの2次情報が蓄積されている。

【0011】データ処理装置11は利用者間類似度算出手段111と適合度推定手段112と項目選択手段113とによって構成され、利用者間類似度算出手段111は入力装置1を介して要求者からの要求が入力されると、項目評価情報記憶部121に蓄積された項目評価情報を使用して他の利用者と要求者の類似度を算出する。また、適合度推定手段112は項目評価情報記憶部121に蓄積された項目評価情報と利用者間類似度算出手段111で算出した利用者間の類似度情報とから、要求者のまだ見聞きしていない対象項目が要求者の興味や嗜好に合うか否かの度合いを示す適合度を推定する。項目選択手段113は適合度推定手段112が推定した適合度を参照して要求に合う対象項目を選択し、結果を出力装置2に出力する。

【0012】次に、図10を参照して従来の方式の動作について説明する。

【0013】図10において、入力装置1を介して要求が入力されると、まず利用者間類似度算出手段111は項目評価情報記憶部121に蓄積されている項目評価情報を使用して他の利用者と要求者の類似度を算出する（ステップE1）。

【0014】例えば、上記MITのフィルタリングシステムでは利用者間の類似度を算出する算出式としてピアソンの積率相関係数を使用している。

【0015】次に、適合度推定手段112は項目評価情報記憶部121に蓄積された項目評価情報と利用者間類似度算出手段111で算出した利用者間の類似度情報とから、要求者に対する各項目毎の適合度を推定する（ステップE2）。

【0016】なお、上記MITのフィルタリングシステム

ムでは適合度を下記式により推定している。

*【数1】

【0017】

*

$$\text{適合度} = \frac{\sum_k (\text{要求者とk番目の利用者の類似度} \times k \text{番目の利用者の項目評価値})}{\sum_k (\text{要求者とk番目の利用者の類似度})}$$

最後に、項目選択手段113は適合度推定手段112で推定した適合度を使用して要求に合う対象項目を選択し（ステップE3）、結果を出力装置2に出力する（ステップE4）。

【0018】

【発明が解決しようとする課題】しかしながら上記したような従来の情報フィルタリング方式では、ある特定の興味や嗜好に利用者が偏って集まると、それに適合度の推定が影響されるという問題があった。

【0019】例えば、a、b、c、dの4つの対象項目を持ち、利用者が4人のシステムに新しい利用者Aが加わった場合を考える。ここで、利用者Aは対象項目a、b、c全てを「良い」と評価しているものとする。このときフィルタリングシステムは対象項目dが利用者Aの興味に合うか否かを推定する。

【0020】表1は利用者A以外の利用者を、対象項目a、b、cに対する評価で分類し、利用者Aとの類似度、対象項目dの評価の平均、および人数を示したものである。

【0021】

【表1】

aの評価	bの評価	cの評価	Aとの類似度	dの評価	人数
1	1	1	1	1	1
1	-1	1	1/3	1	1
-1	1	1	1/3	-1	1
-1	-1	1	1/3	1	1

なお、ここでは評価結果を、「良い」=1、「悪い」=-1と数値化し、類似度の算出にピアソンの積率相関係数を使用している。

【0022】上記従来の方式にしたがって対象項目dに対する利用者Aの評価を推定すると、推定結果は0.5となる。しかしながら、例えば、a：悪い、b：良い、c：良い、d：悪いと評価する他の利用者の人数が10人に増えた場合、推定結果は-7/13になってしまう。

【0023】このように従来の方式では、同じ評価をする利用者の人数が変わるだけで適合度の推定値が大きく変わっていた。

【0024】また、従来の方式では、精度の高い適合度の推定を行うためには多数の項目評価情報を必要とする。一般に新しい対象項目は項目評価情報の数が少ないため、新しい対象項目ほど適合度の推定精度が悪くなるという問題があった。

【0025】ところで、項目評価情報と共に項目評価情

報以外の情報も合わせて使用し、それら複数の情報から適合度を推定する情報フィルタリング方式の場合、どの種類の情報にどれだけの重みを持たせるかという情報の重み付けを行うことが望ましい。このとき、論文と音楽のように対象項目の分野が異なったり、子供と大人のように利用者層が異なると、フィルタリング精度を最良にする重み付けが異なる場合が多い。その場合、情報の重み付けを変更できない方式にするとフィルタリング精度が悪くなるという問題が発生する。

【0026】本発明は上記したような従来の技術が有する問題点を解決するためになされたものであり、第1の目的は、ある特定の興味や嗜好に利用者が偏って集まった場合でも影響されずに適合度を推定し、フィルタリング精度の高い情報フィルタリング方式を提供する。

【0027】また、第2の目的は、新しい項目などのように評価数の少ない項目についても、精度良くフィルタリングを行うことができる情報フィルタリング方式を提供する。

【0028】さらに、第3の目的は、情報に対する重み付けの変更を可能にして、フィルタリング精度に高く保つことができる情報フィルタリング方式を提供する。

【0029】

【課題を解決するための手段】上記目的を達成するため本発明の情報フィルタリング方式は、多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、推薦要求やフィルタリング要求が入力される入力装置と、情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者との類似度を算出する利用者間類似度算出手段と、前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記利用者間類似度算出手段の算出結果から、似た興味や嗜好を持つ利用者の人数を求め、該人数から各利用者近傍での局所的な人口密度を計算する人口密度算出手段を有し、前記適合度推定手段は、前記項目評価情報、前記類似度、及び前記人口密度から前記適合度を推定することを特徴とする。

【0030】また、多数の情報やサービスの中から利用

者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、推薦要求やフィルタリング要求が入力される入力装置と、情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者その他の利用者との類似度を算出する利用者間類似度算出手段と、前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記記憶装置に、前記項目評価情報とともに、情報やサービスに対する評価を行った時刻の情報を各利用者毎に蓄積し、前記時刻情報から利用者の評価情報に対する信頼度の高さを示すLE度を算出するLE度算出手段を有し、前記適合度推定手段は、前記項目評価情報、前記類似度、及び前記LE度から前記適合度を推定することを特徴とする。

【0031】また、多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、推薦要求やフィルタリング要求が入力される入力装置と、情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、前記項目評価情報を使用して、推薦要求やフィルタリング要求を行った利用者である要求者その他の利用者との類似度を算出する利用者間類似度算出手段と、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する第1の適合度推定手段と、該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記記憶装置に、前記項目評価情報とともに、情報やサービスに対する評価を行った時刻の情報を各利用者毎に蓄積し、前記項目評価情報および前記類似度から、各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す推定適合度を算出する第2の適合度推定手段と、前記推定適合度及び前記時刻情報から利用者の評価情報に対する信頼度の高さを示すLE度を算出するLE度算出手段とを有し、前記第1の適合度推定手段は、前記項目評価情報、前記類似度、及び前記LE度から前記適合度を推定することを特徴とする。

【0032】さらに、多数の情報やサービスの中から利用者の興味や嗜好に合うものを推薦する情報フィルタリング方式であって、推薦要求やフィルタリング要求が入力される入力装置と、情報やサービスに対する複数の利用者の評価結果である項目評価情報が蓄積される記憶装置と、前記項目評価情報を使用して、推薦要求やフィル

タリング要求を行った利用者である要求者その他の利用者との類似度を算出する利用者間類似度算出手段と、前記項目評価情報および前記類似度から、フィルタリング対象である各情報やサービスが前記要求者の興味や嗜好に合うか否かの度合を示す適合度を推定する適合度推定手段と、該適合度から要求に合う情報やサービスを選択して出力する項目選択手段と、前記項目選択手段で選択した情報やサービスを表示する出力装置とを有する情報フィルタリング方式において、前記記憶装置に、適合度を推定する際に使用される各情報に対して、情報やサービスの分野及び利用者層に応じて付与される重み情報が蓄積される重み情報記憶部を有し、前記適合度推定手段は、前記項目評価情報、前記類似度、および前記重み情報から前記適合度を推定することを特徴とする。

【0033】このとき、前記重み情報は前記入力装置から入力されてもよい。

【0034】上記のように構成された情報フィルタリング方式は、人口密度算出手段によって似た興味や嗜好を持つ利用者の人数が求められ、その人数から利用者近傍での局所的な人口密度が計算される。

【0035】人口密度は利用者が偏って集まっている付近では高く、利用者が少ないところでは低くなるため、特定の興味や嗜好に利用者が集まる度合いを示している。例えば、利用者間の類似度を人口密度で割ると、適合度を推定する際の項目評価情報の重みが、利用者が偏って集まっている項目では減り、少ない項目では増える。これによって、利用者が特定の興味や嗜好に偏って集まる影響が取り除かれる。

【0036】また、LE度算出手段によって利用者の評価情報に対する信頼度の高さを示すLE度が算出されことで、例えばLE度を要求者と似た評価をより早くするほど大きくなるように設定すれば、要求者と似た興味や嗜好を持ち、かつ、他の利用者よりも早く評価を行う信頼度の高い利用者が、適合度の計算で重要視されるようになる。したがって、項目評価情報が少ない新しい項目についても、利用者に対する適合度を高い精度で推定することができる。

【0037】また、重み情報記憶部に情報やサービスの分野及び利用者層に応じて付与される重み情報を記憶すること、あるいは入力装置からそのような重み情報を入力することで、情報やサービスの分野または利用者層に合わせて、要求者に対する適合度を推定するのに最適な重みを指定することができる。

【0038】

【発明の実施の形態】次に、本発明について図面を参照して説明する。

【0039】（第1実施例）図1は本発明の情報フィルタリング方式の第1実施例の構成を示すブロック図であり、図2は図1に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【0040】図1において、本実施例の情報フィルタリング方式は、入力装置1と、フィルタリングを要求した利用者（以下要求者と称す）に対する各対象項目毎の適合度を推定するデータ処理装置3と、複数の利用者による項目評価情報が蓄積される記憶装置4と、出力装置2とによって構成されている。なお、入力装置1および出力装置2は従来と同様の構成である。

【0041】記憶装置4は項目評価情報記憶部41と項目情報記憶部42とによって構成され、項目評価情報記憶部41には複数の利用者による項目評価情報が蓄積され、項目情報記憶部42には対象項目を識別するためのID番号や項目名などの2次情報が蓄積されている。

【0042】また、データ処理装置3は利用者間類似度算出手段31と人口密度算出手段34と適合度推定手段32と項目選択手段33とによって構成されている。

【0043】利用者間類似度算出手段31は入力装置1を介して要求者からの要求が入力されると、項目評価情報記憶部41に蓄積された項目評価情報を使用して他の利用者と要求者の類似度を算出する。

【0044】人口密度算出手段34は利用者間類似度算出手段31で算出された利用者間の類似度情報から、似た興味や嗜好を持った利用者が何人位居るかという各利用者近傍での局所的な人口密度を算出する。

【0045】適合度推定手段32は項目評価情報記憶部41に蓄積された項目評価情報と利用者間類似度算出手段31で算出された利用者間類似度情報と人口密度算出手段34で算出された局所的な人口密度情報とから要求者に対する各対象項目毎の適合度を推定する。

【0046】項目選択手段33は適合度推定手段32で推定した適合度を参照して、要求に合った対象項目を選択し、結果を出力装置2に出力する。

【0047】次に、図2を参照して本実施例の情報フィルタリング方式の動作について説明する。

【0048】入力装置1を介して推要求が入力される *

$$\text{適合度} = \frac{\sum_k \frac{\text{要求者とk番目の利用者の類似度} \times k \text{番目の利用者の項目評価値}}{k \text{番目の利用者の付近での局所人口密度}}}{\sum_k \frac{\text{要求者とk番目の利用者の類似度}}{k \text{番目の利用者の付近での局所人口密度}}}$$

このように、同じ評価をした利用者が増えて評価結果に偏りが生じても適合度の推定結果に対する影響が少なくなっている。

【0054】人口密度は利用者が偏って集まっている付近では高く、利用者が少ないところでは低くなるため、特定の興味や嗜好に利用者が集まる度合いを示している。そこで、利用者間の類似度を人口密度で割ると、適合度を推定する際の項目評価情報の重みが、利用者が偏って集まっている対象項目では減り、少ない対象項目では増える。したがって、利用者が特定の興味や嗜好に偏って集まる影響が取り除かれ、フィルタリング精度の高い情報フィルタリング方式を得ることができる。（第2

＊と、まず、利用者間類似度算出手段31は項目評価情報記憶部41に蓄積されている項目評価情報を使用して要求者との類似度を算出する（ステップA1）。類似度の算出方法としては、例えば「多変量解析ハンドブック」現代数学社1986年の224頁から228頁に、5種類の距離および6種類の類似度算出方法が掲載されている。これらの算出方法の中から対象項目の性質などを考慮して適当なものを選択して行う。次に、利用者間類似度算出手段31で算出された利用者間の類似度をもとに人口密度算出手段34で各利用者近傍での局所的な人口密度を算出する（ステップA2）。

【0049】続いて、適合度推定手段32は、項目評価情報記憶部41に蓄積されている項目評価情報と、利用者間類似度算出手段31で算出した利用者間類似度情報と、人口密度算出手段34で算出された局所的な人口密度情報とから、要求者に対する各対象項目毎の適合度を推定する（ステップA3）。

【0050】最後に、項目選択手段33は適合度推定手段32で推定した適合度を参照して、要求に合った対象項目を選択し（ステップA4）、結果を出力装置2に出力する（ステップA5）。

【0051】ここで、例えば表1に示した数値例を使って適合度を具体的に算出すると、従来の情報フィルタリング方式では、a：悪い、b：良い、c：良い、d：悪いと評価する人数が1人から10人に増えた場合に、利用者Aに対する対象項目dの適合度の算出結果は、0.5から7/13に変動していた。

【0052】本実施例の情報フィルタリング方式によって利用者Aに対する対象項目dの適合度を推定すると0.5となる。ただし、同じ評価をした人数から人口密度を算出し、下記式により推定評価を求める。

【0053】

【数2】

実施例）次に、本発明の情報フィルタリング方式の第2実施例について図面を参照して説明する。

【0055】図3は本発明の情報フィルタリング方式の第2実施例の構成を示すブロック図であり、図4は図3に示した情報フィルタリング方式の動作手順を示すフローチャートである図3において、本実施例の情報フィルタリング方式は、入力装置1と、各対象項目毎に要求者に対する適合度を推定するデータ処理装置5と、複数の利用者による項目評価情報が蓄積される記憶装置6と、出力装置2とによって構成されている。なお、入力装置1及び出力装置2は従来と同様の構成である。

【0056】記憶装置6は項目評価情報記憶部61と項

目情報記憶部62とによって構成され、項目評価情報記憶部61にはあらかじめ複数の利用者による項目評価情報とその評価結果が入力された時刻(評価時刻)とが蓄積され、項目情報記憶部62にはあらかじめ対象項目を識別するためのID番号や項目名などの2次情報が蓄積されている。

【0057】また、データ処理装置5は利用者間類似度算出手段51とLE度算出手段55と適合度推定手段52と項目選択手段53とによって構成されている。

【0058】利用者間類似度算出手段51は入力装置1を介して要求者からの要求が入力されると、項目評価情報記憶部61に蓄積された項目評価情報を使用して他の利用者と要求者の類似度を算出する。

【0059】LE度算出手段55は項目評価情報記憶部61に蓄積されている項目評価情報と評価時刻とから、利用者の評価情報に対する信頼度の高さを示すリーディングエッジ度(LE度)を各々の利用者毎に算出する。

【0060】適合度推定手段52は、利用者間類似度算出手段51で算出された利用者間の類似度情報と、LE度算出手段55で算出された各利用者のLE度と、項目評価情報記憶部61に蓄積されている項目評価情報とか*

$$Lab = \frac{\sum_{k=ab \text{ ともに評価した項目}} \left(\text{評価の類似度} \times 2 - \frac{k \text{ を評価した利用者数}(k \text{ が最初に評価された時刻} - \text{利用者が} k \text{ を評価した時刻})}{k \text{ が最初に評価された時刻} - \text{現在の時刻}} \right)}{ab \text{ ともに評価した項目の数}}$$

なお、類似度としては評価の差の絶対値の逆数や-(評価の差の絶対値)などを使用すると良い。

【0067】続いて、適合度推定手段52は、利用者間類似度算出手段51で算出された利用者間の類似度情報と、項目評価情報記憶部61に蓄積されている項目評価情報と、LE度算出手段55で算出された各利用者のLE*

$$\text{適合度} = \frac{\sum_k k \text{ 番目の利用者のLE度} \times \text{要求者と} k \text{ 番目の利用者の類似度} \times k \text{ 番目の利用者の項目評価値}}{\sum_k k \text{ 番目の利用者のLE度} \times \text{要求者と} k \text{ 番目の利用者の類似度}}$$

最後に、項目選択手段53は適合度推定手段52で推定した適合度を参照して、要求に合った対象項目を選択し(ステップB4)、結果を出力装置2に出力する(ステップB5)。

【0069】以上説明したように、LE度は、利用者の評価情報に対する信頼度の高さを示すように、要求者と似た評価をより早くするほど大きくなるようにしているため、要求者と似た興味や嗜好を持ち、かつ、他の利用者よりも早く評価を行う信頼度の高い利用者が、適合度の計算で重要視されるようになる。したがって、新しい対象項目についても要求者に対する適合度を従来よりも高い精度で推定することができる。

【0070】(第3実施例)次に、本発明の第3実施例について図面を参照して説明する。

【0071】図5は本発明の情報フィルタリング方式の第3実施例の構成を示すブロック図であり、図6は図5

*ら、要求者に対する各対象項目毎の適合度を推定する。

【0061】項目選択手段53は適合度推定手段52で推定した適合度を参照して、要求に合った対象項目を選択し、結果を出力装置2に出力する。

【0062】次に、図4を参照して本実施例の情報フィルタリング方式の動作について説明する。

【0063】入力装置1から要求が入力されると、まず、利用者間類似度算出手段51は項目評価情報記憶部61に蓄積されている項目評価情報を利用して利用者間の類似度を算出する(ステップB1)。

【0064】次に、LE度算出手段55は項目評価情報記憶部61に蓄積されている項目評価情報と評価時刻とから各利用者のLE度を算出する(ステップB2)。

【0065】LE度は、利用者の評価情報に対する信頼度の高さを示すように、要求者と似た評価をする利用者のうち、より早く評価を行った利用者ほど大きな値を付与するように設定する。ここで、要求者aからみた他の利用者bのLE度Labは例えば下記式で計算することができる。

【0066】
【数3】

※E度とから、要求者に対する各対象項目毎の適合度を推定する(ステップB3)。LE度を導入した適合度は例えば下記式で推定することができる。

【0068】
【数4】

に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【0072】本実施例の情報フィルタリング方式は、上記した第2実施例のデータ処理装置の構成に加え、第2の適合度推定手段76を有する点で異なっている。

【0073】図5において、本実施例の情報フィルタリング方式は、入力装置1と、各対象項目毎に要求者に対する適合度を推定するデータ処理装置7と、複数の利用者による項目評価情報が蓄積される記憶装置8と、出力装置2とによって構成されている。なお、入力装置1及び出力装置2は従来と同様の構成である。

【0074】記憶装置8は項目評価情報記憶部81と項目情報記憶部82とによって構成され、項目評価情報記憶部81は複数の利用者による項目評価情報とその評価結果が入力された時刻とがあらかじめ蓄積され、項目情報記憶部82は対象項目を識別するためのID番号や項

目名などの2次情報があらかじめ蓄積されている。

【0075】また、データ処理装置7は利用者間類似度算出手段71と第1の適合度推定手段72とLE度算出手段75と第2の適合度推定手段62と項目選択手段73とによって構成されている。

【0076】利用者間類似度算出手段71は入力装置1を介して要求者からの要求が入力されると、項目評価情報記憶部81に蓄積された項目評価情報を使用して他の利用者と要求者の類似度を算出する。

【0077】第2の適合度推定手段76は項目評価情報記憶部81に蓄積されている項目評価情報と、利用者間類似度算出手段71で算出された類似度情報とから要求者に対する各対象項目毎の適合度を推定する。

【0078】LE度算出手段75は第2の適合度推定手段76で推定した適合度と項目評価情報記憶部81に蓄積されている評価時刻とから各利用者のLE度を算出する。第1の適合度推定手段72は利用者間類似度算出手段71で算出された利用者間の類似度情報と、LE度算出手段75で算出された各利用者のLE度と、項目評価情報記憶部81に蓄積された項目評価情報とから、要求者に対する各対象項目毎の適合度を推定する。

【0079】項目選択手段73は第1の適合度推定手段*

$$Lab = \frac{\sum_{k=b \text{ が評価した項目}} \left(\text{評価の類似度} \times 2 - \frac{k \text{ を評価した利用者数} (k \text{ が最初に評価された時刻} - \text{利用者が} k \text{ を評価した時刻})}{k \text{ が最初に評価された時刻} - \text{現在の時刻}} \right)}{b \text{ が評価した項目の数}}$$

ここで、各利用者間の類似度は推定適合度と利用者bの評価の差の絶対値の逆数、あるいは－（推定適合度とbの評価の差絶対値）になる。

【0086】続いて、第1の適合度推定手段72は、利用者間類似度算出手段71で算出された利用者間の類似度情報と、項目評価情報記憶部81に蓄積されている項目評価情報と、LE算出手段75で算出された各利用者のLE度とから、要求者に対する各対象項目毎の適合度を推定する（ステップC4）。

【0087】最後に、項目選択手段73は適合度推定手段72で推定した適合度を参照して、要求に合った対象項目を選択し（ステップC5）、結果を出力装置2に出力する（ステップC6）。

【0088】このような構成にすることで、要求者の要求に合うと思われる項目に対して、LE度を用いて適合度を再推定するため、よりフィルタリング精度の高い方式を得ることができる。

【0089】（第4実施例）次に本発明の情報フィルタリング方式の第4実施例について図面を参照して説明する。

【0090】図7は本発明の情報フィルタリング方式の第4実施例の構成を示すブロック図であり、図8は図7に示した情報フィルタリング方式の動作手順を示すフローチャートである。本実施例の情報フィルタリング方式

*72で推定した適合度を参照して、要求に合った対象項目を選択し、結果を出力装置2に出力する。

【0080】このような構成において、次に本発明の第3実施例の動作について図6を参照して説明する。

【0081】なお、図6に示すステップC1およびステップC4～C6の動作は、上記第2実施例と同様である。

【0082】入力装置1から要求が入力されると、まず、利用者間類似度算出手段71は項目評価情報記憶部81に蓄積されている項目評価情報を利用して利用者間の類似度を算出する（ステップC1）。

【0083】次に、第2の適合度推定手段76は第1実施例と同様にして要求者に対する各対象項目毎の適合度を推定する（ステップC2）。

【0084】LE度算出手段75は利用者の項目評価情報のかわりに第2の適合度推定手段76で推定した要求者に対する適合度（以下推定適合度と呼ぶ）を使用してLE度を算出する（ステップC3）。そのため、要求者aからみた他の利用者bのLE度Labは、例えば下記式を用いて算出する。

【0085】

【数5】

は、記憶装置10に情報に対する重みが蓄積される重み情報記憶部103を有し、データ処理装置9に利用者間類似度算出手段で算出される利用者間類似度情報以外の情報を生成する情報生成手段97を有する構成である。

【0091】図7において、本実施例の情報フィルタリング方式は、入力装置1と、対象項目毎に要求者に対する適合度を推定するデータ処理装置9と、複数の利用者による項目評価情報が蓄積される記憶装置10と、出力装置2とによって構成されている。なお、入力装置1及び出力装置2は従来と同様の構成である。

【0092】記憶装置10は項目評価情報記憶部101と項目情報記憶部102と重み情報記憶部103とによって構成されている。

【0093】項目評価情報記憶部101は複数の利用者による項目評価情報等があらかじめ蓄積され、項目情報記憶部102は対象項目を識別するためのID番号や項目名などの2次情報があらかじめ蓄積され、重み情報記憶部103は適合度を推定する際に使用される各情報に対する重み情報があらかじめ蓄積されている。

【0094】また、データ処理装置9は利用者間類似度算出手段91と情報生成手段97と適合度推定手段92と項目選択手段93とによって構成されている。

【0095】利用者間類似度算出手段91は入力装置1を介して要求者からの要求が入力されると、項目評価情

報記憶部101に蓄積された項目評価情報を使用して他の利用者と要求者の類似度を算出する。

【0096】情報生成手段97は、利用者間類似度算出手段91で算出された利用者間類似度情報、項目評価情報記憶部101に蓄積された項目評価情報、および項目情報記憶部102に蓄積された項目情報の中から必要な情報を使用して利用者間類似度情報以外の情報を算出する。

【0097】適合度推定手段92は、項目評価情報記憶部101に蓄積された項目評価情報、利用者間類似度算出手段91で算出された利用者間類似度情報、重み情報記憶部103に蓄積された重み情報、および情報生成手段97で算出された情報の複数の情報から要求者に対する各対象項目毎の適合度を推定する。

【0098】項目選択手段93は適合度推定手段92で推定した適合度を参照して、要求に合った対象項目を選択し、結果を出力装置2に出力する。

【0099】このような構成において、次に本実施例の情報フィルタリング方式の動作について図8を参照して説明する。

【0100】なお、図8に示したステップD1およびステップD4、D5の動作は従来と同様である。

【0101】入力装置1から要求が入力されると、まず、利用者間類似度算出手段91は項目評価情報記憶部101に蓄積されている項目評価情報を利用して利用者*

$$\text{適合度} = \frac{\sum_k \text{正規化された要求者と}k\text{番目の利用者の類似度} \times \text{正規化された情報生成手段からの情報} \times k\text{番目の利用者の項目評価値}}{\sum_k \text{正規化された要求者と}k\text{番目の利用者の類似度} \times \text{正規化された情報生成手段からの情報}}$$

したがって、対象項目の分野や利用者層に合わせて、要求者に対する適合度を推定するのに最適な重みを指定することができるため、フィルタリング精度を向上させることができる。

【0107】なお、重み情報は、重み情報記憶部103に蓄積される必要はなく、入力装置1から入力されても本実施例と同様の効果を得ることができる。

【0108】

【発明の効果】本発明は以上説明したように構成されているので、以下に記載する効果を奏する。

【0109】請求項1に記載のものにおいては、利用者が特定の興味や嗜好に偏って集まる影響が取り除かれ、フィルタリング精度の高い情報フィルタリング方式を得ることができる。

【0110】請求項2及び3に記載のものにおいては、項目評価情報が少ない新しい情報やサービス等についても、利用者に対する適合度を高い精度で推定することができるため、フィルタリング精度を向上させることができる。

【0111】請求項4及び5に記載のものにおいては、情報やサービスの分野あるいは利用者層に合わせて、要求者に対する適合度を推定するのに最適な重みを指定す

*間の類似度を算出する(ステップD1)。

【0102】次に、情報生成手段97は利用者間類似度算出手段91で算出された利用者間類似度情報、項目評価情報記憶部101に蓄積されている項目評価情報、および項目情報記憶部102に蓄積されている項目情報のうち、必要な情報を使用して適合度生成手段92に提供する情報を生成する(ステップD2)。

【0103】続いて、適合度推定手段92は、利用者間類似度算出手段91で算出された利用者間の類似度情報と、項目評価情報記憶部101に蓄積されている項目評価情報と、情報生成手段97で算出された情報と、重み情報記憶部103に蓄積された対象項目毎の重み情報とから、要求者に対する各対象項目毎の適合度を推定する(ステップD4)。

【0104】最後に、項目選択手段93は適合度推定手段92で推定した適合度を参照して、要求に合った対象項目を選択し(ステップD5)、結果を出力装置2に出力する(ステップD6)。

【0105】ここで、適合度推定手段92では、重み情報にしたがって、それぞれの情報の標準偏差が指定された重みになるように正規化が行われ、正規化された値で適合度が推定される。このとき適合度は下記式によって推定される。

【0106】

【数6】

ることができるため、フィルタリング精度の高い情報フィルタリング方式を得ることができる。

【図面の簡単な説明】

【図1】本発明の情報フィルタリング方式の第1実施例の構成を示すブロック図である。

【図2】図1に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【図3】本発明の情報フィルタリング方式の第2実施例の構成を示すブロック図である。

【図4】図3に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【図5】本発明の情報フィルタリング方式の第3実施例の構成を示すブロック図である。

【図6】図5に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【図7】本発明の情報フィルタリング方式の第4実施例の構成を示すブロック図である。

【図8】図7に示した情報フィルタリング方式の動作手順を示すフローチャートである。

【図9】従来の情報フィルタリング方式の構成を示すブロック図である。

【図10】図9に示した情報フィルタリング方式の動作

手順を示すフローチャートである。

【符号の説明】

1 入力装置

2 出力装置

3、5、7、9 データ処理装置

4、6、8、10 記憶装置

31、51、71、91 利用者間類似度算出手段

32、52、92 適合度推定手段

33、53、73、93 項目選択手段

* 34 人口密度算出手段

41、61、81、101 項目評価情報記憶部

42、62、82、102 項目情報記憶部

55、75 LE度算出手段

72 第1の適合度推定手段

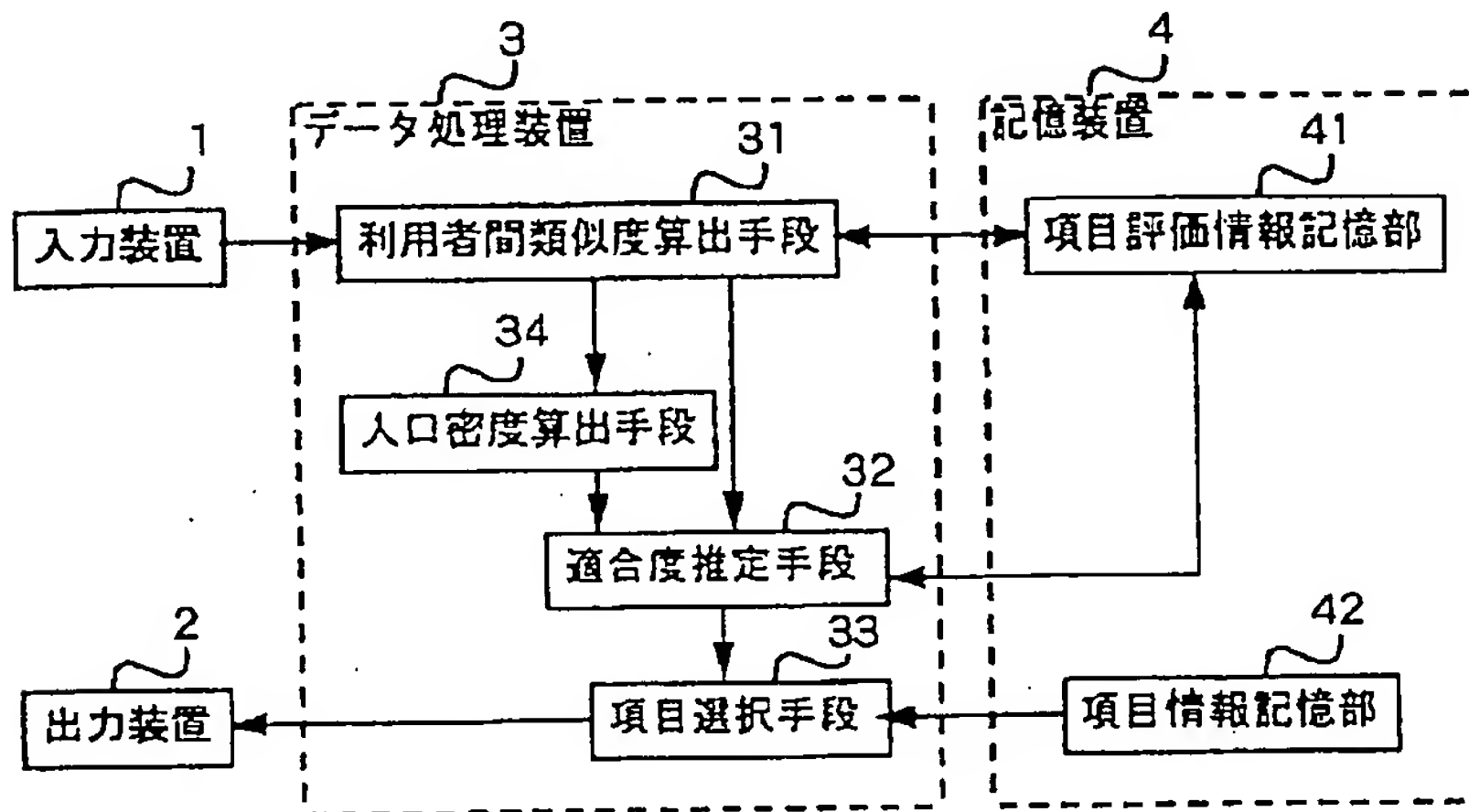
76 第2の適合度推定手段

97 情報生成手段

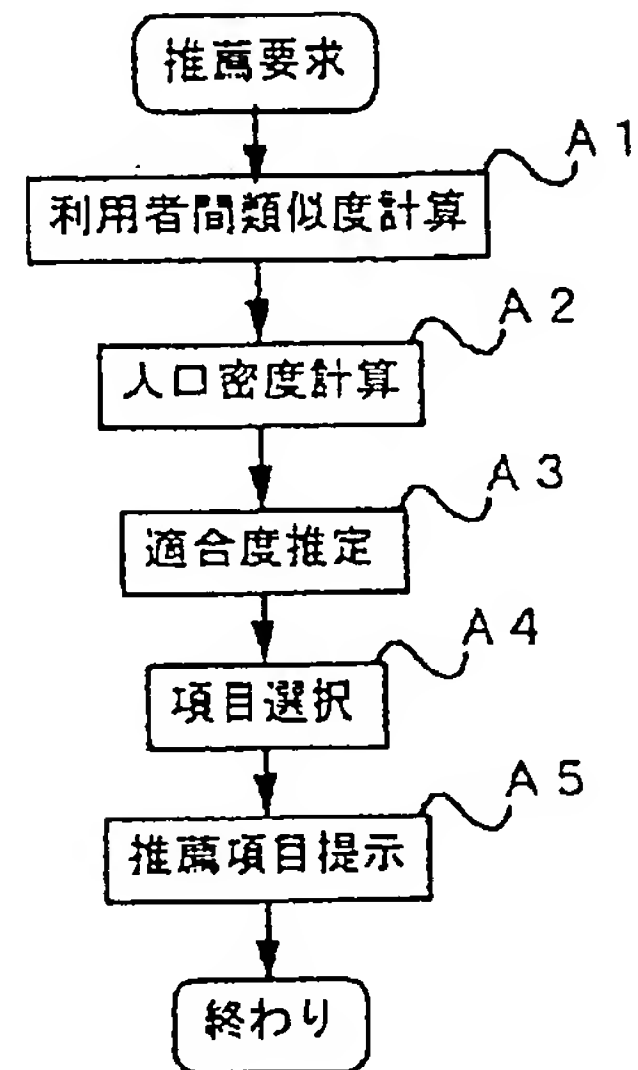
103 重み情報記憶部

*

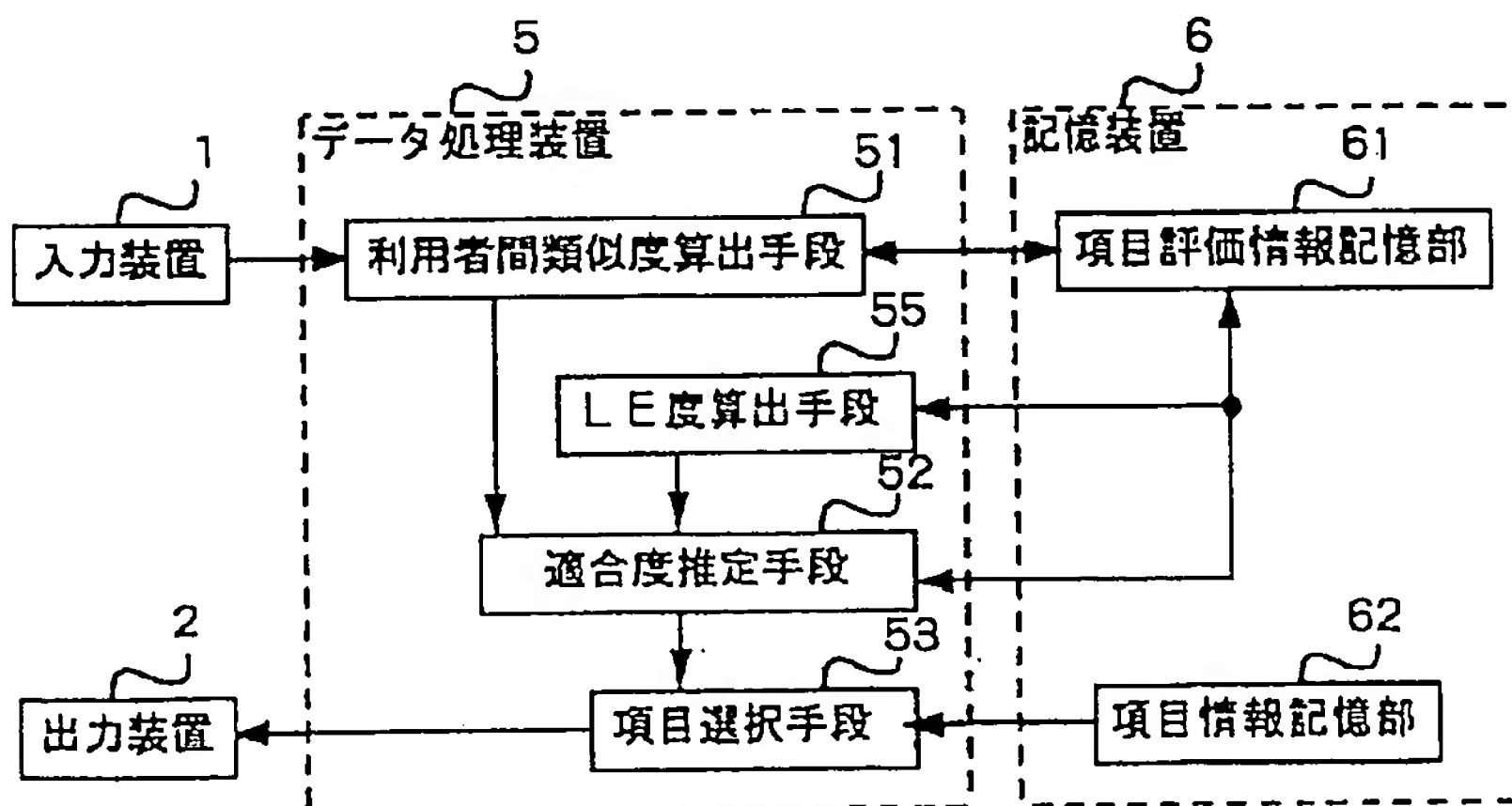
【図1】



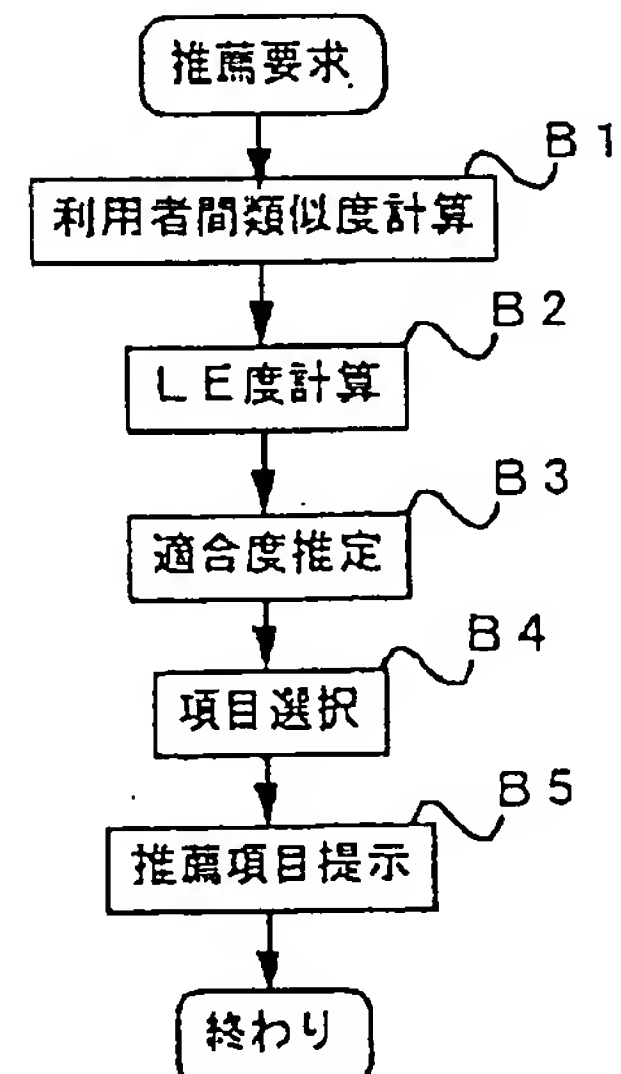
【図2】



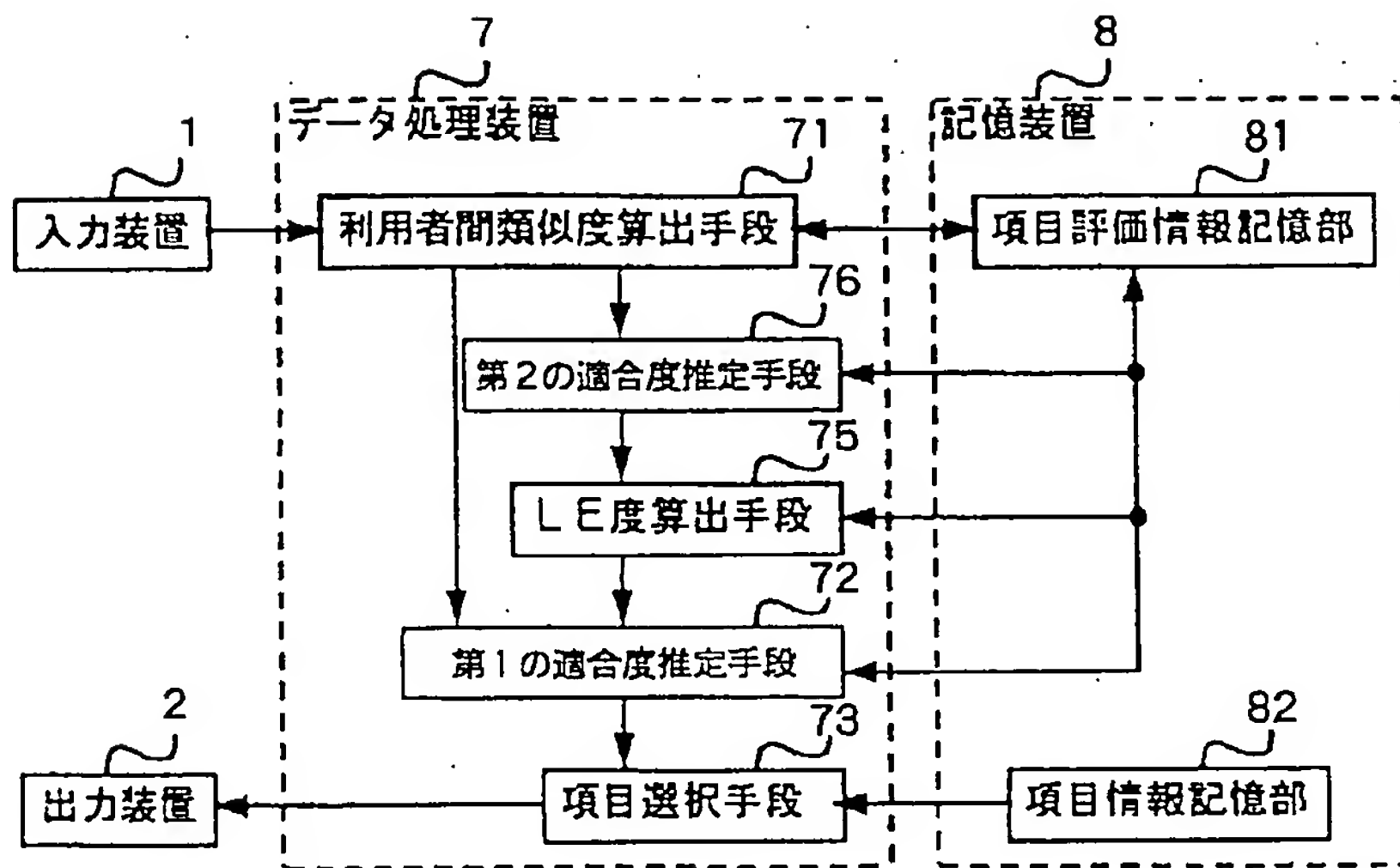
【図3】



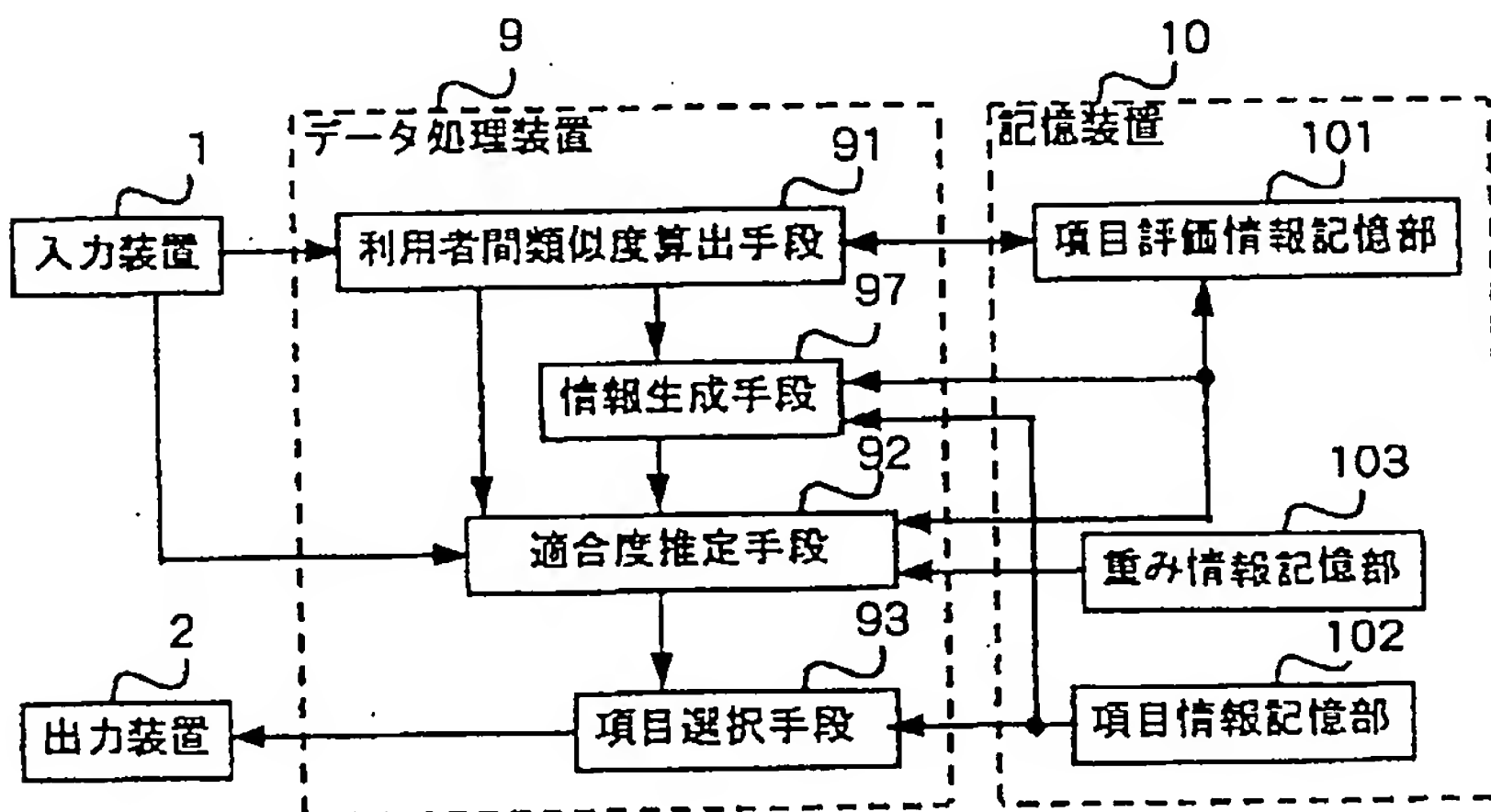
【図4】



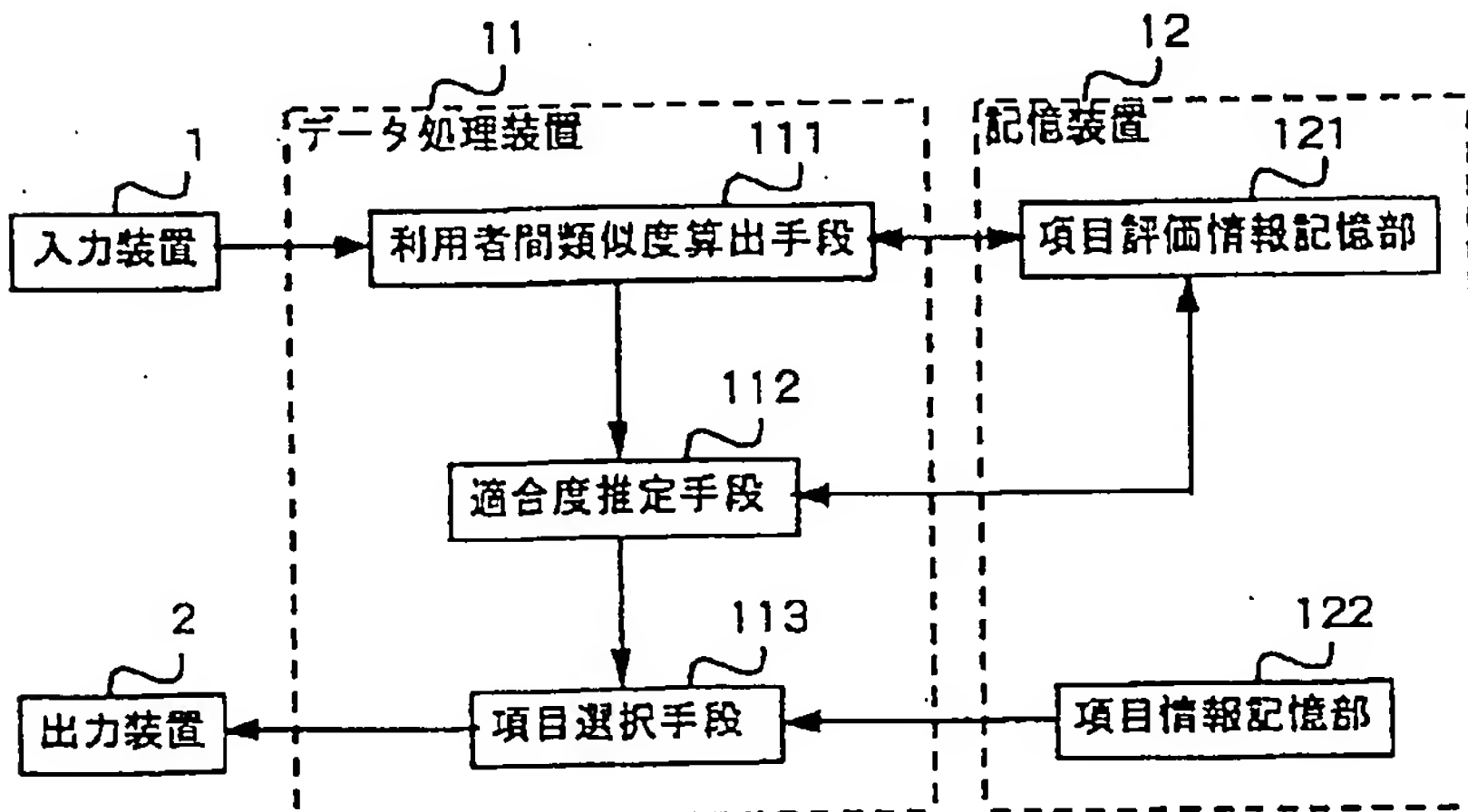
【図5】



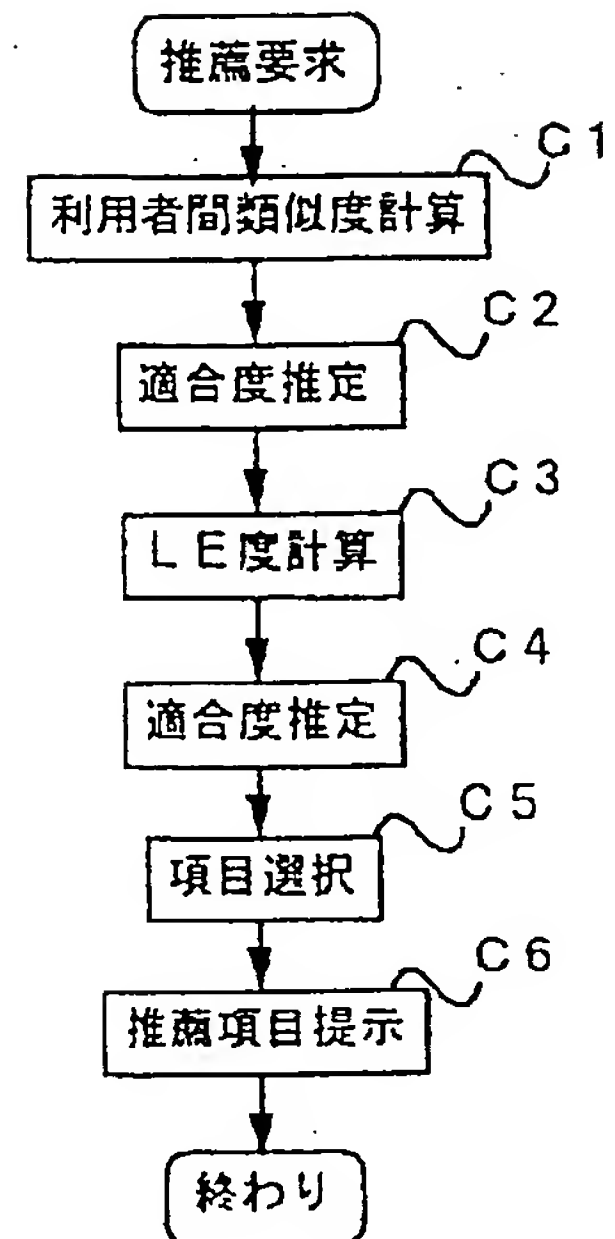
【図7】



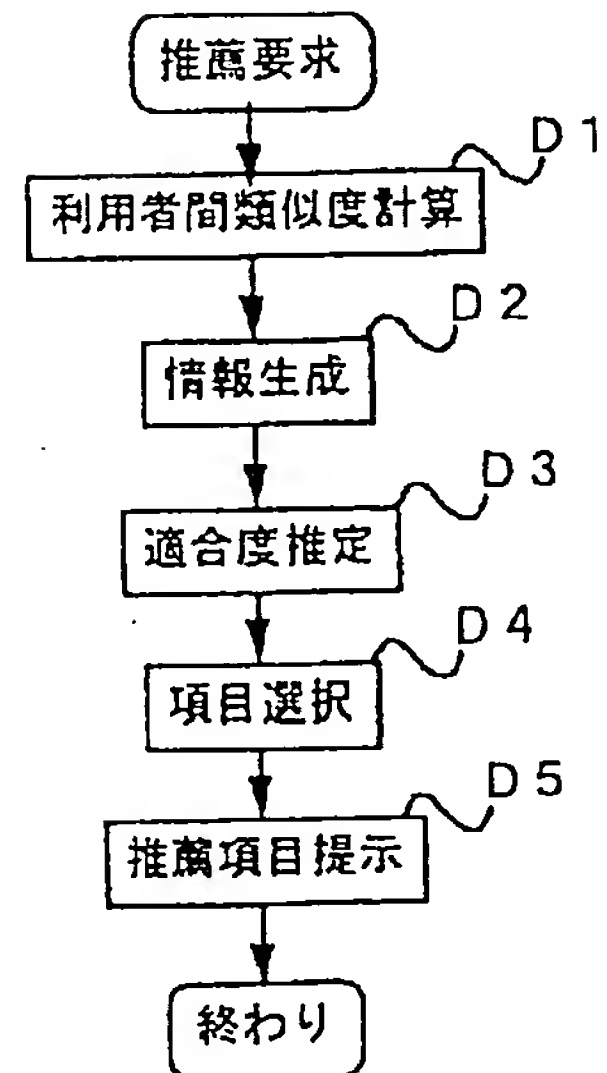
【図9】



【図6】



【図8】



【図10】

